# Classification of Alzheimer's disease subjects from MRI using hippocampal visual features

Olfa Ben Ahmed, Jenny Benois-Pineau, Michèle Allard, Chokri Ben-Amar, Gwénalle Catheline

**Noname manuscript No.**
(will be inserted by the editor)

# Classification of Alzheimer's disease subjects from MRI using hippocampal visual features

**Olfa Ben Ahmed · Jenny Benois-Pineau · Michèle Allard · Chokri Ben Amar · Gwénaëlle Catheline · for the Alzheimer's Disease Neuroimaging Initiative ***

Received: date / Accepted: date

**Abstract** Indexing and classification tools for Content Based Visual Information Retrieval (CBVIR) have been penetrating the universe of medical image analysis. They have been recently investigated for Alzheimer's disease (AD) diagnosis. This is a normal "knowledge diffusion" process, when methodologies developed for multimedia mining penetrate a new application area. The latter brings its own specificities requiring an adjustment of methodologies on the basis of domain knowledge. In this paper, we develop an automatic classification framework for AD recognition in structural Magnetic Resonance Images (MRI). The main contribution of this work consists in considering visual features from the most involved

*Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

O. Ben Ahmed J. Benois Pineau
University of Bordeaux
E-mail: olfa.ben-ahmed,jenny.benois@labri.fr

C. Ben Amar
University of Sfax
E-mail: chokri.benamar@ieee.org

M. Allard G. Catheline
University of Bordeaux
E-mail: michelle.allard,gwenaelle.catheline@chu-bordeaux.fr

region in AD (hippocampal area) and in using a late fusion to increase precision results. Our approach has been first evaluated on the baseline MR images of 218 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database and then tested on a 3T weighted contrast MRI obtained from a subsample of a large French epidemiological study : "Bordeaux dataset". The experimental results show that our classification of patients with AD versus NC (Normal Control) subjects achieves the accuracies of 87% and 85% for ADNI subset and "Bordeaux dataset" respectively. For the most challenging group of subjects with the Mild Cognitive Impairment (MCI), we reach accuracies of 78.22% and 72.23% for MCI versus NC and MCI versus AD respectively on ADNI. The late fusion scheme improves classification results by 9% in average for these three categories. Results demonstrate very promising classification performance and simplicity compared to the state-of-the-art volumetric AD diagnosis methods.

**Keywords** Content Based Visual Indexing · Visual features · Circular Harmonic Functions descriptors · SVM · Bag-of-Visual-Words · Late Fusion · Hippocampus · CSF

## 1 Introduction

Due to an enormous increase of the diversity and of the volume of biomedical image collections and the large range of image modalities getting available nowadays, there is a need for providing automated tools to index medical information. Indexing of medical images using their visual content has shown its efficiency relative to textual approaches [24]. Image content analysis and classification methodologies are now more and more used for medical information mining and retrieval [23] [18] with the aim of Computer-Aided Diagnosis (CAD).

With the aging of population in developed countries, the dementia diseases become a major problem of public health. Alzheimer's disease (AD) is one of the most frequent pathologies and its early detection is very important to achieve delay

in the disease progression. Since AD affects brain cells and causes their degeneration, advances and evolution of medical imaging techniques allow for studying structural changes in human brain and their relationship with clinical diagnosis of AD. Medical information from Magnetic Resonance Imaging (MRI) is used for detecting structural abnormalities of the human brain. In particular, structural MRI measurements help in detecting and tracking the evolution of brain atrophy which is considered as a marker of AD process. In AD, the most common pronounced change in the brain structure is the reduction of the volume of the hippocampus [38]. Several works in the literature use extracted features from the hippocampus region of interest (ROI) for the purpose of diagnosis [15] [9] [13] [16]. Most of the recently proposed approaches do not take into account the visual morphological changes in brain regions, which is our goal. Furthermore, the most of proposed methods for AD diagnosis are time consuming. They require a clinician intervention and suffer from ROI segmentation's errors [15][10], as they are built on the basis of a fine image segmentation. However, hippocampus is not sufficient for the separation of subject with MCI and AD. Other features derived from known biomarkers can be of help. Recent studies on AD diagnosis found that the quantity of cerebrospinal fluid (CSF) in hippocampal region is a biomarker of AD [31]. Indeed, smaller hippocampal volume is associated with greater CSF amount. Also, the authors in [40] proved that the combination of CSF amount and MRI biomarkers provides better prediction than either MRI or CSF alone.

The multimodal nature of multimedia data yielded an active research in fusion of heterogeneous data for classification purposes [4]. Nevertheless, an efficient application of image classification methods in Computer-Aided Diagnosis of AD is not straightforward. Indeed, the specific nature of MRI collections vs general purpose image databases requires an in-depth study of the specific features that explain visible and invisible abnormalities for the diagnosis process. Despite MR image comparison techniques have been proposed in this area, visual features extraction methods have not been yet fully exploited to describe such data. Hence,

in this work we propose an automatic content based framework for recognition of Alzheimer's disease using MRI scans. There are 3 different categories of subjects to recognize: Normal Control (NC), Alzheimer's Disease (AD) patients and the most challenging group Mild Cognitive Impairment (MCI). We propose to automatically extract visual information from MRI using the approximation of image signal by Circular Harmonic Functions (CHF) [34]. To effectively represent content information extracted from the hippocampus ROI, we use the Bag-of-Visual-Words (BoVW)approach [12] . We propose a late fusion scheme, where the probabilistic outputs of classifies on both local features and the amount of CSF are fused to perform the final classification of the MRI scans.

The rest of the paper is organized as follows: Section 2 presents the related works. In section 3, we describe the extraction of visual features with its particularities for this kind of data. In section 4 we present the Late Fusion scheme. In section 5, we present experiments and results and the final section concludes the work and outlines its perspectives.

## 2 Related work

MRI classification task plays an important role in medical image retrieval, which is a part of decision making in medical image analysis. It involves grouping MRI scans into pre-defined classes or finding the class to which a subject belongs. There have been several attempts in the literature to automatically classify structural brain MRI as AD, MCI or NC. Among of the most common methods we find Voxel Based Morphometry (VBM)[3] which is an automatic tool. It allows an exploration of the differences in local concentration of gray matter and white matter. Tensor Based Morphometry (TBM) [39] was proposed to identify local structural changes from the gradients of deformations fields. Object Based Morphometry (OBM) [21] was introduced to perform shape analysis of anatomical structures and recently, features based morphometry (FBM) [35] was proposed as a method for relevant brain features comparison using a probabilistic model on local image features in

scale-space. Some other studies focused on measuring morphological structure of a Region of Interest (ROI) known to be affected by AD such as the hippocampus region. In research studies that analyze hippocampus atrophy,we can identify the following image analysis methods. In [9] [11] [19], the authors automatically segment the hippocampus and use its volume for the classification. In addition to volumetric methods, several surface-based shape description approaches have been proposed to understand the development of AD. In [15] [16] [13], shape information in the form of spherical harmonics (SH) has been used as features in the support vector machine (SVM) classifier. In [32], Statistical Shape Models (SSMs) have been used to model the variability in the hippocampal shapes among the population. Hence, the image-based diagnostic of AD relies mainly on analysis of hippocampus. Nevertheless, the overall volumetric or shape analysis of the hippocampus does not describe the local change of its structure, which is helpful for diagnosis. ROI-based methods are time consuming and observer-dependent. Moreover, most of the approaches cited above were proposed for group analysis and cannot be used to classify individual patients. In order to overcome all these limitations, computer vision tools and visual image processing techniques have been developed to allow an automated detection of atrophy in the ROI.

Recently, Content Based Visual Information Indexing methods have been widely used for medical image analysis. But, few are the works that address the visual content of brain scans to extract information relative to AD. In [1], the authors are focusing on integrating different kinds of information, including textual data, image visual features extracted from scans as well as direct user (doctor) input. In general, features can contain coefficients of a spectral transform of image signal, e.g. Fourier or Discrete Cosine Transform coefficients (DCT), statistics on image gradients [29] [14], etc. Features used in [1] to describe brain images are local binary patterns (LBP) and DCT. [2] uses visual image similarity to help early diagnosis of Alzheimer. [2] [28] prove the performance of user feedback for brain image classification. In the related work [22], Circular Harmonic Functions and

Scale Invariant Features Transform (SIFT)[20] descriptors are computed around the hippocampus region. Then several modes of classification are used to compare images. In [37], the authors propose a ROI retrieval method for brain MRI, they use the LBP and Kanade-Lucas-Tomasi features to extract local structural information. Some works [14] [29] on MRI classification for AD diagnosis evaluate the suitability of the Bag-of-Visual-Words (BoVW) approach for automatic classification of MR images in the case of Alzheimer's disease. In [14], the authors use SIFT descriptors extracted from the whole subject's brain to classify between brain with and without AD. In [29], the authors show that the Bag-of-Features (BOF) approach is able to describe the visual information for discriminating healthy brains from those suffering from the AD. However, both works do not address the MCI case which has become an important construct in the study of AD. The BoVW model represents a whole brain scan or a ROI as a histogram of occurrence of quantized visual features, which are called visual words. The latter received the name of visual signature of an image/ROI.

The choice of the initial description space (features) is of a primary importance as it has to be adapted to the nature of the images. Indeed, despite the good performances of SIFT features reported in [29], there is still place for an intensive investigation of the descriptors choice. SIFT or their approximated version SURF[5], widely used in classification of general purpose image data sets, are not optimal for MRI with the lack of pronounced high frequency texture and clear structural models. Here, based on the work of [22] [6], we resort to Circular Harmonic Functions (CHF) which give interesting approximations of blurred and noisy signal as we have it in MRI. As shown in [34], these descriptors in some cases are superior to SIFT which is a current benchmark. Furthermore, these features as computed on patches inside the ROI or selected on the whole brain, convey local structural information of image signal.

Pattern recognition techniques are widely used in the context of AD diagnosis and in particular Support Vector Machines (SVM) classifiers have proven to be

efficient. High recognition rates are achievable (SVM) focused on brain ROI [15][13] [16] [17]. In this paper, we propose a pattern recognition framework to detect Alzheimer's disease from structural MR images. We extract only visual features from the hippocampal region to emphasize the difference or similarity of subjects with respect to AD. Two kinds of features are extracted : visual local descriptors using the Circulars Harmonic Functions and the amount of CSF pixels in the hippocampal area. These features are of different nature. Hence, it is appropriate to deploy the multimedia fusion approaches, such as reported in [4] despite we are working with the same imaging modality.

## 3 Visual content description

### 3.1 Extraction of visual features from hippocampal area

Visual features extraction is a common step in the overall processing chain yielding image interpretation and classification. Applied to MRI, it has to be populated by particular techniques already in use for brain MRI analysis. As the visual information has to be extracted from a specific anatomical region, an atlas-based selection of this region has to be performed. In brain image analysis, such an atlas called Automated Anatomical Labeling (AAL) [36] does exist for a normalized "statistical " brain, a so-called "template" MNI [1]. Hence, the first step as depicted in Figure 1, consists in normalizing the brain image to be analyzed with regard to this template.

#### *3.1.1 Image normalization*

The normalization of the available 3D brain data is an adjustment of overall size and orientation to the template MNI. Linear (affine) and non-linear alignment is possible. In our work, we choose to apply an affine registration because we look for preserving a specific pattern of the ROI and avoiding features deformation. Affine

[1] http://www.mni.mcgill.ca/

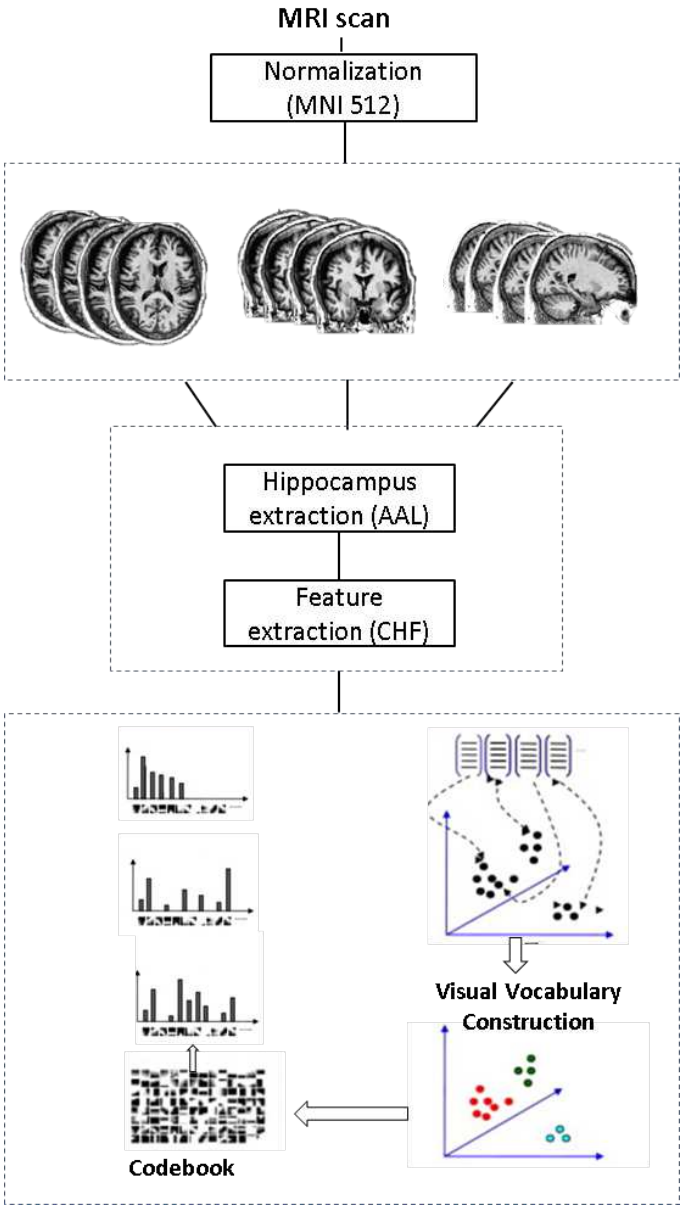


**Fig. 1** Visual feature extraction from the hippocampal ROI

registration is done using the Montreal Neurological Institute template (MNI)512.

Statistical Parametric Mapping (SPM)[2] and the VBM toolbox [3] are used to fulfill the registration.

### *3.1.2 Hippocampus ROI extraction*

To extract the hippocampus ROI, we use the AAL Atlas. The template-aligned 3D MRI of a brain is superimposed on the 3D atlas and only voxels which are labeled in AAL as hippocampal are selected (see the second block in Figure 1). Once the hippocampal region has been roughly delimited, the features can be extracted and signature built.

### *3.1.3 Signature generation*

After brain alignment and ROI's selection, we compute image features. As it was already noted, we need to extract only those features, which contain visual information related to the presence or absence of the AD. In this work, Circular Harmonic Functions (CHF) were used for selection of contrasted patterns in brains as it is the case in [6].

*Circular Harmonic Functions (CHF)*

Gauss-Laguerre Harmonic Functions are complex-valued radial profile functions multiplied by complex exponent:

$$\Psi(r,\theta;\sigma) = \Psi_n^{|\alpha|}(\frac{r^2}{\sigma})e^{i\alpha\theta} \tag{1}$$

Their radial profiles are Laguerre functions:

$$\Psi_n^{|\alpha|}(x) = \frac{1}{\sqrt{n!\Gamma(n+\alpha+1)}}x^{\frac{\alpha}{2}}e^{\frac{-x}{2}}L_n^{\alpha}(x) \tag{2}$$

[2] (Welcome Laboratory of the Department of Cognitive Neurology, Institute of Neurology, London, United Kingdom, http://www.fil. ion.ucl.ac.uk./spm/)

[3] http://dbm.neuro.uni-jena.de/vbm

where $n = 0, 1, ...; \alpha \pm 1, \pm 2...$ and $L_n^{\alpha}(x)$ are Laguerre polynomials. $r$, $\theta$ are polar coordinates, $\sigma$ is a scale parameter and $\Gamma$ is a gamma function.

$$L_n^{\alpha}(x) = (-1)^n x^{-\alpha} \exp^x \frac{d}{dx^n}(x^{n+\alpha} e^{-x}) \tag{3}$$

LG-CHF is complete orthogonal set of functions on the real plane. Thus, the image $I(x, y)$ can be expanded in the analysis point $x_0$ , $y_0$ for fixed scale $\sigma$ in Cartesian system. The coefficients of the partial expansion of local neighborhood can be used as a feature descriptor. The advantages of these features are such that they capture both the direction and smoth variations of image signal. Their draw back is in a rather slow convergence, hence a sufficient number of coefficients has to be retained for image description. The number of coeficients retained define the dimensionnality of the descriptor. The reasonable dimensionality of 150 coefficients (see [6][22] ) was used in the present work. Hence the dimension of the descriptor is comparable with that one of conventional SIFT. More mathematical details about the CHF descriptors can be found in [33] [34].

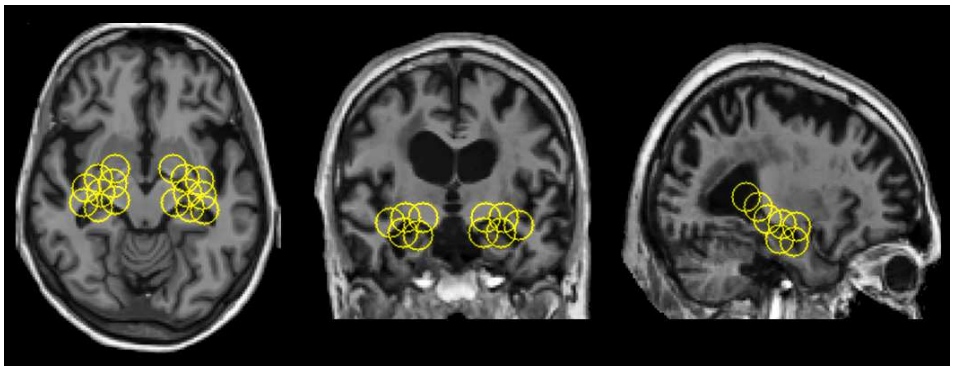

**Fig. 2** Illustrating of feature placement of local CHF features on the hippocampus ROI (axial, coronal and sagittal projections) of an MRI scan from the ADNI dataset

Figure 2 shows an illustration of CHF features extraction on hippocampus. The extracted feature points "support areas" (i.e. where the descriptors are computed) are denoted with yellow circles. Here, the scans are densely sampled in a selected

hippocampal ROI by a grid of circular patches and the signal decomposition on a CHF basis is computed for each patch. We perform selection sampling and 2D CHF transform computation on a slice-by-slice basis. Hence the whole description of the hippocampal volume is a collection of 2D CHF descriptors for each slice and each projection of the selected volume. The CHF coefficients extracted from several areas overlapping with the mask are different and depend on the signal presented in the ROI (atrophy or not).

Recently, Bag-of-Visual-Words (BoVW) approach has shown its power in this field, modeling the hippocampus ROI as a set of local features[6]. The role of BoVW model is to cluster extracted features from hippocampus in order to build a visual vocabulary. The region's shape differs from one projection to another. Thus, we choose to perform clustering 3 times from different projections (sagittal, axial and coronal) and to generate one visual vocabulary per projection. Firstly, all features $f^s_{n,i}$ , here $n$ and $i$ stand respectively for slice and feature indexes, are extracted from the ROI on all slices for the sagittal projection then features are clustered by k-means algorithm. The same is done for axial and coronal projections. All features $f^s_{n,i}$, $f^a_{n,i}$ , $f^c_{n,i}$ and centers of clusters $c_{sk}$ ,$c_{ak}$, $c_{ck}$ obtained by k-means (where K is the codebook size) here have the same dimensionality of the descriptor being used. In case of SIFT it is 128 and for CHF it is 150. According to the BoVW approach, we then call cluster centers "visual words". Once the visual words have been determined, the image signature per projection is generated. Each feature is assigned to closest visual word using the distance $d(f^s_{n,i}, c_s)$ , in our case the Euclidean distance is used. Then each projection is represented by a normalized histogram of occurrence of visual words. The image signature $h$ is built by concatenating the histograms from all projections $h = [h_s h_a h_c]$.

### 3.2 CSF volume computation

The increased quantity of CSF in the hippocampal region is an important visual biomarker for AD diagnosis. Indeed in the case of AD, the hippocampus shrinks

and the liberated volume is filled with CSF. To analyze the shrinkage, we count the CSF voxels in the region of the hippocampus. In the MRI T1 scans the CSF is appearing as dark areas, thus we can select it just by thresholding.

$$B^*(x, y, z) < T_{dark}$$

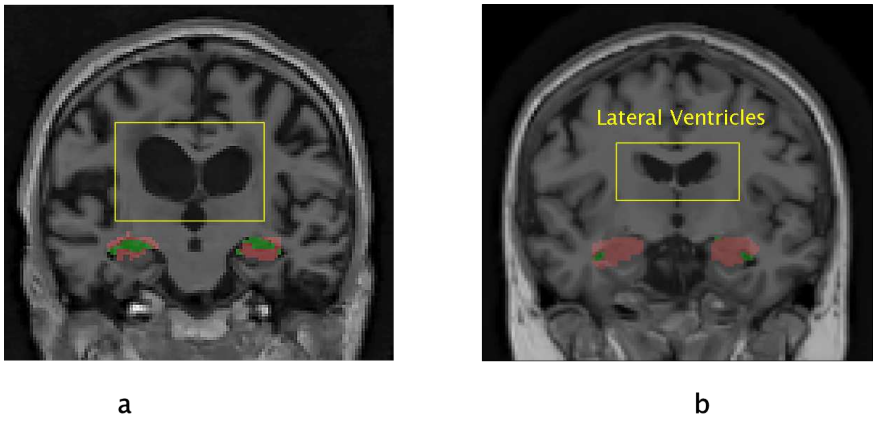


**Fig. 3** CSF on the hippocampus region : a) AD brain, b) Healthy : Coronal projection brain

But due to the large difference in brightness and contrast of MRI scans, all scans need to be transformed in such a way that, similar intensities will have similar tissue-specific meaning. In our work, we perform the scale normalization method proposed in [25]. In order to select the optimal threshold, the following procedure is performed: all voxels from the hippocampus regions from all scans are collected together and the threshold between dark (hyposignal) and bright (hyper signal) voxels is estimated using Otsu method [26]. In fact, one threshold for all images is computed. But normal patients have a little CSF amount in the hippocampus area. Thus, to ensure correct delineation when computing the threshold by Otsu method, we add additional regions where CSF is always present: The Lateral Ventricles (LV) by referring to the domain knowledge. In addition, adding some pixels from the Lateral Ventricles may improve the discrimination results because AD patients show more CSF in the LV than do MCI and NC subject. Using this

procedure, the volume of the CSF in a normalized hippocampal area is measured in a quantity of voxels. It will be later denoted by $V$. Figure 3 illustrates the results of detection of CSF (in green) in hippocampal region. The CSF (green color) is situated around the hippocampus (red color) boundaries. The added region is marked with a yellow rectangle. It can be seen that the quantity of CSF in the case of AD (a) is higher. Hence the quantity of CSF in the hippocampus area is determined for all scans in the database and Bayesian classifier is trained using this parameter to distinguish the subjects.

## 4 Classification approach

The classification approach we design is aimed to combine the two sources of information : visual signature and the volume of CSF in a global decision framework to discriminate between AD and NC, MCI and NC and AD and MCI subjects. Taking into account the advances in multimedia fusion research in the literature, we propose to do it by a late fusion scheme. The overall diagram of the approach is presented in Figure 4. The CHF-based visual signatures are first classified between the categories two by two with a state-of-the art SVM approach with an Radial Basis Function (RBF) kernel. The classification of subjects on the basis of the CSF volume is performed by a Bayesian classifier. Indeed, we have here a scalar feature and the class probabilities can reasonably be a priori trained (AD are much more rare in patients cohorts, than NC and MCI for instance). Both classification schemes give a decision output. We transform it into an homogeneous probabilistic output and form the second order feature vectors of dimension 2. They are then submitted to the trained SVM binary classifier for each classification problem given above (AD vs NC, NC vs MCI and AD vs MCI). We stress that in this work we address a binary classification problem as the goal is to assess the discriminative power of our scheme which uses both hippocampal shape expressed by CHF features and CSF volume biomarkers in an automatic classification of cohorts. We now present the details of each step of the approach.

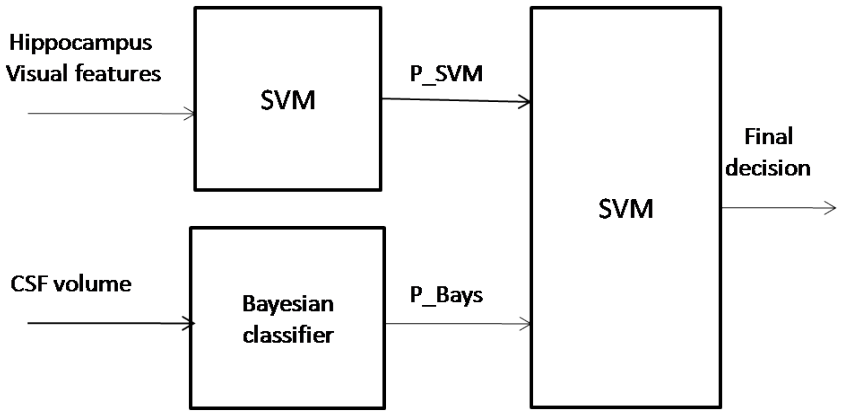


**Fig. 4** Late Fusion scheme

### 4.1 Support vector machines classifier for visual signatures classification

Hence, we solve a set of binary classification problems AD vs NC, NC vs MCI and AD vs MCI. The unknown subject is classified by maximizing the score of these three classifiers. We use the well known SVM classifier [7] in the visual description space of signatures built with CHF descriptors. At training step it separates a given set of training data of instance label pairs $(x_i, y_i), i = 1, ..., l$ where $x_i \in R^n$ and $y \in \{1, -1\}$ by maximizing the distances to the hyperplane that separates the two classes in a kernel-transformed space. Then the classification of unknown data is performed in this space accordingly to their position with regard to the hyperplane. For more details on SVMs we refer the reader to [30]. In this work we use the RBF kernel defined by: $exp(-\gamma * |u - v|^2)$. In many settings, for a given input sample and for a given classifier we are more interested in the degree of confidence that the output should be +1. In such cases it is useful to produce a probability $P(y = 1|x)$. Given k classes, for any x, the goal is to estimate

$$p_i = P(y = i|x), i = 1....k$$

We first estimate pairwise class probabilities by Platt approximation [27]

$$r_i = \frac{1}{1+e^{Af_i+B}}$$

where $f$ is the decision value at $x$. $A$ and $B$ are estimated by minimizing the negative log likelihood of training data (using their labels and decision values). For each binary classifier we will have the probabilistic output $P_{SVM_i}(x)$.

### 4.2 Bayesian classifier for CSF based features Classification

The Bayesian classifier uses the parametric model of pdf for each class which we suppose Gaussian. It gives the most likely class for a given observation. Let $V$ denote the CSF volume for a given subject, $Y$ is the subject class label ($Y = AD, NCorMCI$), and $C = 2$ (binary classification) is the number of classes. The problem consists in classifying the sample $v$ to the class $c_*$ maximizing $P(Y = c|V = v)$ over $c = 1, , C$. Applying Bayes rule gives:

$$P(Y = c|V = v) = \frac{P(V=v|Y=c)P(Y=c)}{P(V=v)}$$

and reduces the original problem to:

$$c_* = argmax_{c=1,...C} P(V = v|Y = c)P(Y = c)$$

we denote the related probability of a sample by $P_{Bayes_i}(x)$

### 4.3 SVM-based late fusion scheme

The probabilistic outputs of SVM-based signature classification and CSF volume-based Bayesian classification are now available for each training sample in all binary classification problems . We form the 2 dimensional feature vectors

$$Z(x) = (P_{SVM_i}(x), P_{Bayes_i}(x))^T$$

and the training set pairs $Z(x), y(x)$. Finally, the second SVM classifier in cascade is trained with a linear kernel using Leave-One-Out Cross-Validation.

Metrics used to evaluate final late fusion classification performance are :

- Accuracy (Acc) $= \frac{(TP + TN)}{(TP + TN + FN + FP)}$

- Sensitivity (Sen) $= \dfrac{TP}{(TP + FN)}$
- Specificity (Spe) $= \dfrac{TN}{(TN + FP)}$

Here $TP$ (True Positives) are AD patients correctly identified as AD, $TN$ (True Negatives) are controls correctly classified as controls, $FN$ (False Negatives) are AD patients incorrectly identified as controls and $FP$ (False Positives) are controls incorrectly identified as AD. Similar definition is hold for other binary classification problems NC vs MCI and AD vs MCI.

## 5 Experiments and results

### 5.1 MRI data

Data used in the preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a 60 million, 5- year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimers disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-

2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

In this work, we selected from the ADNI dataset the same data as [40], with the same demographic information for each of the diagnosis groups (NC, AD and MCI). The data sample consists of 218 structural MRIs from the ADNI dataset with 35 Alzheimer's Disease (AD) patients, 72 cognitively normal (NC) and 111 Mild Cognitive Impairment (MCI) subjects. Images are standard 1.5 T screening baseline T1 weighted obtained using volumetric 3D MPRAGE protocol.

The second source of data is a study of AD on a real cohort, we call it "Bordeaux dataset" [4] [8] comprising 37 structural MRI (16 AD and 21 NC).

## 5.2 Results and Discussion

### *5.2.1 CSF volume computation*

In this subsection we give the figures showing the credibility of CSF quantity biomarker extracted with our method (see 3.2). Table 1 presents the quantities of CSF voxels within the hippocampal ROI. Indeed, from Table 1, one can see that the amount of CSF increases from NC to AD.

**Table 1** CSF amounts

| Class | Volume ( mean ± SD ) |
|---|---|
| NC | 212 ± 101 |
| MCI | 316 ± 104 |
| AD | 402 ± 126 |

[4] http://www.incia.u-bordeaux1.fr/

*5.2.2 Classification results*

Table 2 presents classification performance. It summarizes classification results of AD versus NC, MCI versus NC and MCI versus AD for the ADNI subset. We also present classification results of AD versus NC obtained on the "Bordeaux dataset". Since the latter does not contain MCI cases, relative classification problems are not addressed in our experiments.

**Table 2** Classification results

| | | ADNI dataset | | | Bordeaux dataset |
|---|---|---|---|---|---|
| | | AD vs NC | NC vs MCI | MCI vs AD | AD vs NC |
| Hippo VF (CHF) | Acc | 85.05% | 74.32% | 58.9% | 79% |
| | Spe | 94.45% | 85.59% | 63.97 % | 80% |
| | Sens | 65.72% | 57% | 42.86 % | 70% |
| Hippo VF (SIFT) | Acc | 79.44 % | 72.14% | 52.05% | 67.56% |
| | Spe | 93 % | 84.69% | 55.85 % | 76.19 % |
| | Sens | 51.43% | 52.78% | 40 % | 56.25 % |
| Hippo VF (SURF) | Acc | 81.3% | 73.23% | 50% | 66.67 % |
| | Spe | 91.67% | 83.78 % | 57.6% | 40% |
| | Sens | 60% | 56.95% | 25.73% | 85.71% |
| CSF volume | Acc | 78.5 % | 58.47 % | 62.33% | 80% |
| | Spe | 72.29% | 48% | 67.39 % | 72% |
| | Sens | 70.5% | 66% | 60% | 60% |
| Hippo VF (CHF)+ CSF | Acc | 87% | 78.22% | 72.23% | 85% |
| | Spe | 100 % | 83.34% | 70 % | 81% |
| | Sens | 75.5 % | 70.73% | 75 % | 76% |

Firstly, we compare the performance on CHF visual features (3.1.3) with regard to conventional SIFT and SURF ones. It can be seen that the proposed CHF features systematically outperform SIFT and SURF in all three quality metrics: Accuracy, Specificity and Sensitivity. We note that the SURF features with the lowest dimension (64) between three classes of descriptors are not applicable in our

problem. In fact, they are less precise than SIFT and give a very low sensitivity (25.73%) in a difficult case of MCI vs AD. The CHF descriptors used are of a comparable dimension (150) with SIFT (128), but outperform them. The results presented in the Table 2 for visual features alone, correspond to the optimal sizes of visual vocabularies we estimated experimentally optimizing the accuracy criterion. For SIFT features, the size of visual dictionary per projection was 100 yielding to the dimension of 3*100 of the BoVW. For SURF features, it was of 150 yielding the signature size of 3x150. Finally, for the CHF features, the dictionary consisted of 150 visual words yielding respectively the dimension of the visual signature of 3x 150. The low cardinality of the optimal codebook can be explained by a reasonably limited number of descriptors. Indeed, the dense sampling is performed only on the hippocampal ROI in a limited number of slices (70 for sagittal, 97 for axial and 42 for coronal projections respectively).

Using visual features of the hippocampus on the ADNI subset, we achieved an accuracy of 85.05% and 74.32% respectively for AD versus NC and NC versus MCI classification. However, structural change on hippocampus is not sufficiently accurate on its own to be an absolute diagnostic criterion to separate AD from MCI cases. In the case of MCI versus AD classification, performance drops to 58.9%. We aimed to deal with this challenging category (MCI). To enhance the results, CSF amount measurement was added. The CSF amount classification using Bayesian classifier gives an accuracy of 62.33% and 58.47% for the recognition of the MCI cases respectively from the AD and NC subjects. Moreover, we note that adding supplementary voxels from the Lateral Ventricles helps to boost the performance of CSF delineation and thus improve the classification results. Indeed, the accuracy of AD vs NC classification by CSF amount increases from 74.1% to 78.5%. Hence, we retained this finding for classification and all results in Table 2 were obtained with this approach. Since those two kinds of features were extracted from the same brain (hippocampus), our assumption that they could provide complementary information for classification was correct.

For MCI versus AD classification, using the late fusion, we achieved 72.23% of accuracy compared to 58.5% using only CHF features. For the MCI versus NC classification, accuracy increases from 74.32% to 78.22%. As we can see from table 2 the sensitivity values of both AD vs MCI and NC vs MCI classification undergoes a significant increase ( from 42.86% to 75% for the MCI vs AD cases for example) when we use the late fusion. These results show that CSF volume improves the classification accuracy by an average of 9% when combined with the visual signatures especially for the MCI cases classification which is the most challenging task due to the strong heterogeneity of this class.

In a second part of experiments, we selected 15 MRI scans of AD ($60 \pm 3$ years old) and 12 aging subject from the Normal control category ( $80 \pm 6$ years old ) of the ADNI dataset. Our approach distinguishes well between Alzheimer's disease and the aging normal control subjects with an accuracy of 85% and sensitivity of 76%. Hence, adding CSF amount not only improved MCI cases classification but also helped to separate old healthy subjects from those suffering from AD.

We compare our work with results obtained in [40]. First, the authors used the volume and the shape of hippocampus to perform subject categorization and second, they added CSF bimomarkers and volume and shape of the lateral ventricles to improve results in the case of AD and MCI recognition. Our content based approach outperforms all achieved results on [40]. For example better classification accuracy was achieved in AD versus MCI and NC versus MCI classification tasks is 69.6% and 72% respectively, which is lower than results obtained in our present work. In the case of AD or MCI categorization we reached better results (accuracy of 72.23%, a 70% of specificity and of sensitivity 75% ) compared to [40] in which the authors obtained only 69.9% of accuracy, 68.6% of specificity and of sensitivity 70.7%. We can conclude that combining visual features of AD biomarkers performs better than using volume or shape. Also, in [40], the authors use the freesurfer software to select region which is a very time consuming (about hours of processing) task contrarily to the atlas mapping used in our work. Therefore,

the ability to efficiently classify MCI and AD patients based on visual features of structural MRI might shed light on the ability to predict the conversion from MCI to AD, which is of clinical interest.

## 6 Conclusion

In this paper we proposed a simple and robust classification approach of MRI scans for Alzheimer's disease diagnosis. The approach is based on visual content description of anatomical structure of a brain region involved in AD ( hippocampal area). We proposed a late fusion of classification results on two biomarkers: hippocampus and CSF. The experiments showed that combining hippocampus features and CSF amount classification gave better accuracy especially when discriminating between AD and MCI than when using either visual features or CSF volume separately for discriminating between AD and MCI than using either visual features extraction or CSF volume computation separately. We also demonstrated that the proposed method provides better classification accuracy compared to other volumetric methods. In the perspective of this work we plan to use multiple ROIs, but also multiple MRI modalities in the established classification framework.

*Acknowledgments.*

This research is supported by the Franco-Tunisian program, the LaBRI, University of Bordeaux 1 and university of Bordeaux 2. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli

Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

## References

1. Agarwal M, Mostafa J (2010) Image retrieval for Alzheimer disease detection. In: Proceedings of the First MICCAI international conference on Medical Content-Based Retrieval for Clinical Decision Support, Springer-Verlag, Berlin, Heidelberg, MCBR-CDS'09, pp 49–60
2. Akgül CB, Ünay D, Ekin A (2009) Automated diagnosis of Alzheimer's disease using image similarity and user feedback. In: Proceedings of the ACM International Conference on Image and Video Retrieval, ACM, New York, NY, USA, CIVR '09, pp 1–8
3. Ashburner J, Friston KJ (2000) Voxel-Based Morphometry-the methods. Neuroimage 11(6):805–821
4. Ayache S, Quénot G, Gensel J (2007) Classifier fusion for SVM based multimedia semantic indexing. In: Proceedings of the 29th European conference on IR research, Springer-Verlag, Berlin, Heidelberg, ECIR'07, pp 494–504

5. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-Up Robust Features (SURF). Comput Vis Image Underst 110(3):346–359
6. Ben Ahmed O, Benois-Pineau J, Ben Amar C, Allard M, Catheline G (2013) Early Alzheimer disease detection with bag-of-visual-words and hybrid fusion on structural MRI. In: 11th International Workshop on Content-Based Multimedia Indexing (CBMI) 2013, IEEE, pp 79–83
7. Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: the fifth annual workshop on Computational learning theory, ACM, pp 144–152
8. Catheline G, Periot O, et al (2010) Distinctive alterations of the cingulum bundle during aging and Alzheimer's disease. Neurobiology of Aging 31(9):1582 – 1592
9. Chupin M, Gérardin E, Cuingnet R, et al (2009) Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19(6):579–587
10. Chupin M, Hammers A, Liu RSN, et al (2009) Automatic segmentation of the hippocampus and the amygdala driven by hybrid constraints: Method and validation. NeuroImage 46(3):749–761
11. Colliot O, Chételat G, Chupin M, et al (2008) Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. Radiology 248(1):194–201
12. Csurka G, Dance CR, Fan L, et al (2004) Visual categorization with bags of keypoints. In: Workshop on Statistical Learning in Computer Vision, ECCV, pp 1–22
13. Cuingnet R, Gerardin E, Tessieras J, et al (2011) Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage 56(2):766–781
14. Daliri MR (2012) Automated Diagnosis of Alzheimer disease using the Scale-Invariant Feature Transforms in Magnetic Resonance Images. J Med Syst

36(2):995–1000

15. Gerardin E, Chételat G, Chupin M, et al (2009) Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. NeuroImage 47(4):1476–1486
16. Gutman B, Morra YWJ, Toga A, Thompson P (2009) Disease classification with hippocampal shape invariants. Hippocampus 19(6):572–578
17. Klöppel S, Stonnington CM, Chu, et al (2008) Automatic classification of MR scans in Alzheimer's disease. Brain 131(3):681–689
18. Kumar A, Kim J, Cai W, Fulham M, Feng D (2013) Content-based medical image retrieval: A survey of applications to multidimensional and multimodality data. Journal of Digital Imaging 26(6):1–15
19. Liu Y, Paajanen T, Zhang Y, Westman E, et al (2011) Combination analysis of neuropsychological tests and structural MRI measures in differentiating AD, MCI and control groups-the add neuromed study. Neurobiol Aging 32(7):1198–1206
20. Lowe DG (2004) Distinctive Image Features from Scale-Invariant Keypoints. Int J Comput Vision 60(2):91–110
21. Mangin JF, Rivière D, Cachia A, Papadopoulos-Orfanos D, et al (2003) Object-based strategy for morphometry of the cerebral cortex. In: IPMI, Ambleside, UK, LNCS-2732, SpringerVerlag, UK, SpringerVerlag, pp 160–171
22. Mizotin M, Benois-Pineau J, Allard M, Catheline G (2012) Feature-based brain MRI retrieval for Alzheimer disease diagnosis. In: 19th IEEE International Conference on Image Processing (ICIP), pp 1241–1244
23. Müller H, Deserno TM (2011) Content-based medical image retrieval. In: Biomedical Image Processing - Methods and Applications, Springer, pp 471–494
24. Müller H, Michoux N, Bandon D, Geissbuhler A (2004) A review of content-based image retrieval systems in medical applicationsclinical benefits and future directions. International Journal of Medical Informatics 73(1):1–23

25. Nyúl LG, Udupa JK, Zhang X (2000) New variants of a method of mri scale standardization. IEEE Trans Med Imaging 19(2):143–150
26. Otsu N (1979) A Threshold Selection Method from Gray-level Histograms. IEEE Transactions on Systems, Man and Cybernetics 9(1):62–66
27. Platt JC (1999) Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Advances in large margin and classifiers, MIT Press, pp 61–74
28. Ridha BH, Barnes J, Van de Pol LA, et al (2007) Application of automated medial temporal lobe atrophy scale to Alzheimer disease. Archives of Neurology 64(6):849–854
29. Rueda A, Arevalo JE, Cruz-Roa A, Romero E, González FA (2012) Bag of features for automatic classification of Alzheimer's disease in Magnetic Resonance Images. In: CIARP, pp 559–566
30. Scholkopf B, Smola AJ (2001) Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA, USA
31. Shaw LM, Vanderstichele H, Knapik-Czajka M, et al (2009) Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Annals of Neurology 65(4):403–413
32. Shen K, Bourgeat P, Fripp J, Meriaudeau F, Salvado O (2012) Detecting hippocampal shape changes in Alzheimer's disease using statistical shape models. NeuroImage 59(3):2155–2166
33. Sorgi L, Cimminiello N, Neri A (2006) Keypoints Selection in the Gauss Laguerre Transformed Domain. In: BMVC, British Machine Vision Association, pp 539–547
34. Sorokin DV, Mizotin M, Krylov AS (2011) Gauss-laguerre keypoints extraction using fast hermite projection method. In: Proceedings of the 8th international conference on Image analysis and recognition - Volume Part I, Springer-Verlag, Berlin, Heidelberg, ICIAR'11, pp 284–293

35. Toews M, Wells W, Collins DL, Arbel T (2010) Feature-Based Morphometry: Discovering Group-related Anatomical Patterns. NeuroImage 49(3):2318–2327
36. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al (2002) Automated Anatomical Labeling of Activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15(1):273 – 289
37. Ünay D, Ekin A, Jasinschi RS (2010) Local structure-based region-of-interest retrieval in brain MR images. IEEE Transactions on Information Technology in Biomedicine 14(4):897–903
38. Villain N, Desgranges B, Viader F, et al (2008) Relationships between hippocampal atrophy, white matter disruption, and gray matter hypometabolism in Alzheimer's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 28(24):6174–6181
39. Wolz R, Julkunen V, Koikkalainen J, Niskanen Eea (2011) Multi-method analysis of mri images in early diagnostics of alzheimer's disease. PLoS ONE 6(10):e25,446, DOI 10.1371/journal.pone.0025446
40. Yang X, Tan MZ, Qiu A (2012) CSF and Brain Structural Imaging Markers of the Alzheimer's Pathological Cascade. PLoS ONE 7(12):e47,406, DOI 10.1371/journal.pone.0047406